Accounting for model error in variational data assimilation

A.S. Lawless

School of Mathematical and Physical Sciences
University of Reading

Work supported by NERC
Introduction

2 Modifying the observation error covariance

3 Weak-constraint 4D-Var

4 Summary
The 4D-Var data assimilation problem can be expressed as the minimization of

$$
\mathcal{J}[x_0] = \frac{1}{2}(x_0 - x^b)^T B^{-1}(x_0 - x^b) + \frac{1}{2} \sum_{i=0}^{N} (H_i[x_i] - y_i)^T R_i^{-1}(H_i[x_i] - y_i)
$$

subject to the dynamical system

$$
x_{i+1} = M_i(x_i)
$$

where

- $x^b$ A priori (background) estimate
- $y_i$ Observation
- $B$ Background error covariance matrix
- $R_i$ Observation error covariance matrix
- $H_i$ Observation operator
Incremental 4D-Var (Gauss-Newton)

We usually solve the 4D-Var problem by a series of linear quadratic minimizations of the form

\[
\tilde{J}^{(k)}[\delta x_0^{(k)}] = \frac{1}{2} \left( \delta x_0^{(k)} - [x^b - x_0^{(k)}] \right)^T B^{-1} \left( \delta x_0^{(k)} - [x^b - x_0^{(k)}] \right) \\
+ \frac{1}{2} \sum_{i=0}^{N} (H_i \delta x_i^{(k)} - d_i^{(k)})^T R_i^{-1} (H_i \delta x_i^{(k)} - d_i^{(k)})
\]

with

\[
d_i = y_i - H_i [x_i^{(k)}] \\
\delta x_{i+1} = M_i \delta x_i
\]

This is equivalent to a Gauss-Newton iteration (Lawless et al. (2005), Quart. J. Roy. Met. Soc.).
Alternative notation

For the linear case we can write the cost function as follows:

$$J(x_0) = \frac{1}{2}(x_0 - x^b)^T B^{-1}(x_0 - x^b) + \frac{1}{2}(\hat{y} - \hat{H}x_0)^T \hat{R}^{-1}(\hat{y} - \hat{H}x_0),$$

$$\hat{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ \vdots \\ y_N \end{pmatrix}, \quad \hat{H} = \begin{pmatrix} H_0 \\ H_1 \rightarrow 1 \\ \vdots \\ \vdots \\ H_N \rightarrow N \end{pmatrix}, \quad \text{and} \quad \hat{R} = \begin{pmatrix} R_0 & 0 & \cdots & \cdots & 0 \\ 0 & R_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & R_N \end{pmatrix}.$$

A.S. Lawless, a.s.lawless@reading.ac.uk
In practice models contain errors, due to
- Inaccurate parameter specifications
- Inaccurate parametrisations of sub-grid physical processes
- Inaccurate specification of boundary conditions
- Numerical schemes only approximate solutions
- Poor model resolution

We wish to account for this in the data assimilation process.
Model uncertainty

Assumptions

- Linear model
- Additive model error
- Model error unbiased, Gaussian, random
Model uncertainty

Assumptions

- Linear model
- Additive model error
- Model error unbiased, Gaussian, random

True model

\[ x_{i+1}^t = M_i^t x_i^t \]
Model uncertainty

**Assumptions**
- Linear model
- Additive model error
- Model error unbiased, Gaussian, random

True model

\[ x^t_{i+1} = M^t_i x^t_i \]

Erroneous model

\[ x^t_{i+1} = M^t_i x^t_i + \eta_i, \quad \eta_i \sim \mathcal{N}(0, Q_i) \]
Consider again the form of the objective function

$$J(x_0) = \frac{1}{2}(x_0 - x^b)^T B^{-1}(x_0 - x^b) + \frac{1}{2}(\hat{y} - \hat{H}x_0)^T R^*_t^{-1}(\hat{y} - \hat{H}x_0),$$

where now $M$ is the erroneous model and $R^*_t$ is the covariance of $\epsilon_{ob}^* = \hat{y} - \hat{H}x_0$. 
Modifying the observation error covariance (work with Kat Howes and Alison Fowler)

Consider again the form of the objective function

\[ J(x_0) = \frac{1}{2}(x_0 - x^b)^T B^{-1}(x_0 - x^b) + \frac{1}{2}(\hat{y} - \hat{H}x_0)^T R^*^{-1}(\hat{y} - \hat{H}x_0), \]

where now \( M \) is the erroneous model and \( R^* \) is the covariance of \( \epsilon_{ob}^* = \hat{y} - \hat{H}x_0^t \).

How should we specify \( R^* \)?
Let,

\[ R^*_{(i,k)} = \langle \epsilon^*_{obi} (\epsilon^*_{obk})^T \rangle. \]

Then,

\[ R^*_{(i,k)} = \begin{cases} 
R_0 & \text{for } i=k=0 \\
R_i + H_i \left[ \sum_{j=1}^{\min(i,k)} M_{j\rightarrow i} Q_j M_{j\rightarrow k}^T \right] H_k^T & \text{for } i=k \\
H_i \left[ \sum_{j=1}^{\min(i,k)} M_{j\rightarrow i} Q_j M_{j\rightarrow k}^T \right] H_k^T & \text{otherwise.} 
\]
Combined model error and observation error covariance matrix

\[
R^* = \begin{pmatrix}
R_0 & 0 & \cdots & \cdots & 0 \\
0 & R_1 + Q^*(1,1) & Q^*(1,2) & \cdots & Q^*(1,N) \\
\vdots & \vdots & R_2 + Q^*(2,2) & \vdots & \vdots \\
0 & Q^*(N,1) & \cdots & \cdots & R_N + Q^*(N,N)
\end{pmatrix}
\]

- increase in block diagonal terms due to model error;
- off diagonal block model error covariance terms (time correlations of model error);
- model error covariance terms vary over time.
Combined model error and observation error covariance matrix

\[
R^* = \begin{pmatrix}
R_0 & 0 & \cdots & 0 \\
0 & R_1 + Q^*(1,1) & \cdots & 0 \\
0 & Q^*(2,1) & R_2 + Q^*(2,2) & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
0 & Q^*(N,1) & \cdots & \cdots & R_N + Q^*(N,N)
\end{pmatrix}
\]

- increase in block diagonal terms due to model error;
- off diagonal block model error covariance terms (time correlations of model error);
- model error covariance terms vary over time.

How can we calculate this?
Calculating the modified covariance matrix

Define

\[(d^o_b)_i = y_i - H_i M_{0\rightarrow i} x^b.\]
Calculating the modified covariance matrix

Define

$$(d^o_b)_i = y_i - H_i M_{0 \rightarrow i} x^b.$$ 

Then we can show that

$$E[(d^o_b)_i (d^o_b)_k^T] = R^*_{(i,k)} + H_i M_{0 \rightarrow i} B M_{0 \rightarrow k}^T H_k^T.$$ 

Or

$$R^*_{(i,k)} = E[(d^o_b)_i (d^o_b)_k^T] - H_i M_{0 \rightarrow i} B M_{0 \rightarrow k}^T H_k^T.$$
Idealized coupled nonlinear model

Couples the Lorenz 63 system and 2 linear equations*,

\[
\begin{align*}
\dot{x} & = -\sigma x + \sigma y + \alpha v, \\
\dot{y} & = -xz + rx - y + \alpha w, \\
\dot{z} & = xy - bz, \\
\dot{w} & = -\Omega v - k(w - w^*) - \alpha y, \\
\dot{v} & = \Omega(w - w^*) - kv - \alpha x,
\end{align*}
\]

(2)

where \(\sigma = 10, \ r = 30, \ b = \frac{8}{3}, \ k = 0.1, \ \Omega = \frac{\pi}{10}\) and \(w^* = 2\).

- Runge-Kutta 2nd order method with fixed time step \(\Delta t = 0.01\) used to approximate solution of coupled ODE’s.
- Consider this as ‘true’ model.

True idealized coupled nonlinear model

Parameter perturbation method: Stochastic forcing simulation*, but with Gaussian error distributions and random error at each time-step.

The true parameter values $\sigma^t$, $k^t$ and $\alpha^t$ change at every time-step,

- $\sigma_i^t = \gamma_\sigma \sigma$, where $\gamma_\sigma \sim \mathcal{N}(1, \frac{12}{1})$,
- $k_i^t = \gamma_k k$, where $\gamma_k \sim \mathcal{N}(1, \frac{12}{6})$,
- $\alpha_i^t = \gamma_\alpha \alpha$, where $\gamma_\alpha \sim \mathcal{N}(1, \frac{12}{1})$,

The difference between the true and erroneous model at each time can be considered as additive model error $\eta_i$ of the form,

$$x_i^t = M_{i-1}(x_{i-1}^t) + \eta_i \quad i = 1, 2, \ldots 500.$$

Numerical experiments: design

- Assimilation window length 500 time-steps of length $\Delta t = 0.01$, with all variables observed every 10 time-steps directly. Let $B = R_i = 10^{-4}I$.

- Perturb the true model states using $B$ and $R_i$ respectively to produce background model state $x^b$ and observations $y_i$.

- Select background vector $x^b$ and perturb using $B$ to obtain a sample of 20 background values (note these are all at initial time $t_0$).

- For each observation time $t_i$: select observation vector $y_i$ and perturb using $R_i$ to obtain a sample of 20 observations.

- Use these samples to estimate $(d^o_b)_i = y_i - H_i M_{0 \rightarrow i} x^b$ at each observation time $t_i$.

- Take the expectations of the innovation products $E[(d^o_b)_i (d^o_b)_i^T]$ at each observation time $t_i$.

- Calculate $R^*_i = E[(d^o_b)_i (d^o_b)_i^T] - H_i M_{0 \rightarrow i} B M_{0 \rightarrow i}^T H_i^T$.

Compare 4DVar analysis accuracy using $R^*$ as opposed to $\hat{R}$. 

A.S.Lawless, a.s.lawless@reading.ac.uk

Model error in variational assimilation
Method 1: use $\hat{R}$ in 4DVar
Method 2: use $R^*$ in 4DVar

<table>
<thead>
<tr>
<th>Variable</th>
<th>Truth</th>
<th>Analysis Method 1</th>
<th>Error % Method 1</th>
<th>Analysis Method 2</th>
<th>Error % Method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>-3.4866</td>
<td>-3.1111</td>
<td>10.77</td>
<td>-3.4829</td>
<td>0.11</td>
</tr>
<tr>
<td>$y$</td>
<td>-5.7699</td>
<td>-5.2994</td>
<td>8.15</td>
<td>-5.7843</td>
<td>0.25</td>
</tr>
<tr>
<td>$z$</td>
<td>18.341</td>
<td>18.6500</td>
<td>1.68</td>
<td>18.3464</td>
<td>0.03</td>
</tr>
<tr>
<td>$w$</td>
<td>-10.7175</td>
<td>-10.8140</td>
<td>0.90</td>
<td>-10.7181</td>
<td>0.01</td>
</tr>
<tr>
<td>$v$</td>
<td>-7.1902</td>
<td>-7.9787</td>
<td>10.97</td>
<td>-7.1928</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table: Analysis from Method 1 and Method 2.
Numerical experiments: results

Figure: Trajectory for $w$ over the assimilation window.
Weak-constraint 4D-Var
(work with Adam El-Said and Nancy Nichols)

Previously we had the objective function

$$\mathcal{J}[x_0] = \frac{1}{2}(x_0 - x^b)^T B^{-1}(x_0 - x^b) + \frac{1}{2} \sum_{i=0}^{N} (H_i[x_i] - y_i)^T R_i^{-1}(H_i[x_i] - y_i)$$

subject to the dynamical system

$$x_{i+1} = M_i(x_i).$$

We now consider the model as a weak constraint

$$x_{i+1} = M_i(x_i) + \eta_i, \quad \eta_i \sim \mathcal{N}(0, Q_i)$$
Weak-constraint 4D-Var

Error formulation

\[ J(x_0, \eta_0, \ldots, \eta_{N-1}) = J_b + J_o + \frac{1}{2} \sum_{i=0}^{N-1} \eta_i^T Q_i^{-1} \eta_i \]

State formulation

\[ J(x_0, x_1, \ldots, x_N) = J_b + J_o + \frac{1}{2} \sum_{i=0}^{N-1} (x_{i+1} - \mathcal{M}_i(x_i))^T Q_i^{-1} (x_{i+1} - \mathcal{M}_i(x_i)) \]
The expected accuracy of the numerical solution and the speed of convergence are both determined by the condition number of the Hessian.

Condition number

\[ \kappa(A) = \| A \| \| A^{-1} \| \]

In the matrix 2-norm, for a symmetric positive definite matrix \( A \), we have

\[ \kappa(A) = \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)} \]
The expected accuracy of the numerical solution and the speed of convergence are both determined by the condition number of the Hessian.

\[ \kappa(A) = \|A\| \|A^{-1}\| \]

In the matrix 2-norm, for a symm. pos. def. matrix \( A \), we have

\[ \kappa(A) = \lambda_{\text{max}}(A)/\lambda_{\text{min}}(A) \]

What are the condition numbers of the Hessians of the two formulations sensitive to?

(extends previous work on strong-constraint case by Haben, Lawless and Nichols)
Hessian - Error formulation: \( S_p = \nabla^2 \mathcal{J}(x_0, \eta_i) \)

\[
S_p = D^{-1} + L^{-T} H^T R^{-1} H L^{-1}
\]

\[
S_p = \begin{pmatrix}
B_0^{-1} \\
& Q_1^{-1} \\
& & \ddots \\
& & & Q_n^{-1}
\end{pmatrix}
+ \begin{pmatrix}
H_0^T \\
& (H_1 M_1)^T \\
& & (H_2 M_2)^T \\
& & & \ddots \\
& & & & \ddots \\
& & & & & (H_n M_n)^T
\end{pmatrix} R^{-1} \begin{pmatrix}
H_0 \\
& H_1 M_1 \\
& & H_1 \\
& & & \ddots \\
& & & & \ddots \\
& & & & & H_n M_n
\end{pmatrix}
\]
Hessian - State formulation: \( S_x = \nabla^2 \mathcal{J}(x_i) \)

\[
S_x = L^T D^{-1} L + H^T R^{-1} H
\]

\[
S_x = \begin{pmatrix}
B_0^{-1} + M_1^T Q_1^{-1} M_1 & -M_1^T Q_1^{-1} \\
-Q_1^{-1} M_1 & Q_1^{-1} + M_2^T Q_2^{-1} M_2 & -M_2^T Q_2^{-1} \\
-Q_2^{-1} M_2 & & & \ddots \\
H_0^T R_0^{-1} H_0 & \cdots & \cdots & \cdots & H_n^T R_n^{-1} H_n
\end{pmatrix}
\]
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:

\[ \mathbf{B}_0 = \sigma^2_b \mathbf{C}_B \quad \text{and} \quad \mathbf{Q}_i = \sigma^2_q \mathbf{C}_Q \quad \text{for all} \quad i = 1, \ldots, N. \]

\[ \mathbf{C}_B \text{ and } \mathbf{C}_Q \text{ are valid auto-covariance matrices (Symmetric pos-def)}. \]

- Observation errors are uncorrelated, \( \mathbf{R}_i = \sigma^2_o \mathbf{I} \) for all \( i = 0, \ldots, N \).

- We take \( p \) regularly spaced, direct observations at every time step giving \( p(N + 1) \) observations in total.

- Observations available are less than the size of the state, \( p(N + 1) < n(N + 1) \).

- We observe at the same positions at each time step.

- The model \( \mathbf{M} \) is a circulant matrix.
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - \( B_0 = \sigma_b^2 C_B \) and \( Q_i = \sigma_q^2 C_Q \) for all \( i = 1, \ldots, N \).
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - $C_B$ and $C_Q$ are valid auto-covariance matrices. (Symmetric pos-def)
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - $C_B$ and $C_Q$ are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\Rightarrow R_i = \sigma_o^2 I$ for all $i = 0, ..., N$. 

A.S. Lawless, a.s.lawless@reading.ac.uk

Model error in variational assimilation
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma^2_b C_B$ and $Q_i = \sigma^2_q C_Q$ for all $i = 1, \ldots, N$.
  - $C_B$ and $C_Q$ are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies R_i = \sigma^2_o I$ for all $i = 0, \ldots, N$.
- We take $p$ regularly spaced, direct observations at every time step giving $p(N + 1)$ observations in total.
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - $C_B$ and $C_Q$ are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.
- We take $p$ regularly spaced, direct observations at every time step giving $p(N + 1)$ observations in total.
  - Observations available are less than the size of the state, $p(N + 1) < n(N + 1)$.
Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:

- \( B_0 = \sigma_b^2 C_B \) and \( Q_i = \sigma_q^2 C_Q \) for all \( i = 1, \ldots, N \).
- \( C_B \) and \( C_Q \) are valid auto-covariance matrices. (Symmetric pos-def)

Observation errors are uncorrelated \( \implies R_i = \sigma_o^2 I \) for all \( i = 0, \ldots, N \).

We take \( p \) regularly spaced, direct observations at every time step giving \( p(N + 1) \) observations in total.

- Observations available are less than the size of the state, \( p(N + 1) < n(N + 1) \).
- We observe at the same positions at each time step.
Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - $C_B$ and $C_Q$ are valid auto-covariance matrices. (Symmetric pos-def)

- Observation errors are uncorrelated $\implies R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.

- We take $p$ regularly spaced, direct observations at every time step giving $p(N + 1)$ observations in total.
  - Observations available are less than the size of the state, $p(N + 1) < n(N + 1)$.
  - We observe at the same positions at each time step.

- The model $M$ is a circulant matrix.
Bounds have been derived on the condition number of the Hessian $S_p$. These bounds indicate:

- Decreasing $\sigma_o^2$ (specifying accurate observations) increases $\kappa(S_p)$. 

A.S.Lawless, a.s.lawless@reading.ac.uk

Model error in variational assimilation
Error formulation - Bounds

Bounds have been derived on the condition number of the Hessian $S_p$. These bounds indicate:

- Decreasing $\sigma_o^2$ (specifying accurate observations) increases $\kappa(S_p)$.
- Longer assimilation windows increase $\kappa(S_p)$.
Bounds have been derived on the condition number of the Hessian $S_p$. These bounds indicate:

- Decreasing $\sigma_o^2$ (specifying accurate observations) increases $\kappa(S_p)$.
- Longer assimilation windows increase $\kappa(S_p)$
- $\kappa(S_p)$ is linearly influenced by $\kappa(D)$. 
Bounds have been derived on the condition number of the Hessian $S_p$. These bounds indicate:

- Decreasing $\sigma_o^2$ (specifying accurate observations) increases $\kappa(S_p)$.
- Longer assimilation windows increase $\kappa(S_p)$.
- $\kappa(S_p)$ is linearly influenced by $\kappa(D)$.
- As the difference between $\sigma_b^2$ and $\sigma_q^2$ increases, so does $\kappa(S_p)$. 
Bounds have been derived on the condition number of the Hessian $S_p$. These bounds indicate:

- Decreasing $\sigma_o^2$ (specifying accurate observations) increases $\kappa(S_p)$.
- Longer assimilation windows increase $\kappa(S_p)$.
- $\kappa(S_p)$ is linearly influenced by $\kappa(D)$.
- As the difference between $\sigma_b^2$ and $\sigma_q^2$ increases, so does $\kappa(S_p)$.
- $\kappa(B_0)$ and $\kappa(Q_i)$ increase as correlation length-scales are increased.
State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $S_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $S_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
- Interesting feature: $\kappa(S_x)$ is *immune* to assimilation window length iff observation operator is *full rank*.
State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $S_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
- Interesting feature: $\kappa(S_x)$ is *immune* to assimilation window length iff observation operator is *full rank*.
- Otherwise: as assimilation window length increases, so does $\kappa(S_x)$. Due to $\lambda_{min}(L^TD^{-1}L) \rightarrow 0$, (second derivative matrix).
Bounds have been derived on the condition number of the Hessian $S_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
- Interesting feature: $\kappa(S_x)$ is *immune* to assimilation window length iff observation operator is *full rank*.
- Otherwise: as assimilation window length increases, so does $\kappa(S_x)$. Due to $\lambda_{min}(L^T D^{-1} L) \rightarrow 0$, (second derivative matrix).
- *Very* sensitive to condition number of $D$, with the addition of increased sensitivity to $\sigma_q$ (in comparison to $S_p$).
Numerical Results: Error formulation

- We demonstrate the bounds by viewing the condition number as a function of correlation length-scale:

\[ \kappa(S_p) \] (blue-surface) and bounds (red-mesh surfaces) as a function of \( L(C_B) \) and \( L(C_Q) \).

**Figure**: \( \kappa(S_p) \) (blue-surface) and bounds (red-mesh surfaces) as a function of \( L(C_B) \) and \( L(C_Q) \).
Numerical Results: State formulation

The $\mathcal{J}(x_i)$ formulation is more sensitive to $\kappa(D)$:

Figure: Surface plot of $\kappa(S_x)$ (blue surface) and bounds (red mesh). Non-vertical axes measure background error correlation length-scale $L(C_B)$ and model error correlation length-scale $L(C_Q)$. 
Numerical Results: State formulation

- We now consider the condition number as a function of assimilation window length and observation density:

Figure: Surface plot of $\kappa(S_x)$ (blue surface). Vertical axis measures condition number. The non-vertical axes measure spatial observation density per assimilation time step and assimilation window length.
Summary

- Allowing for model uncertainty becoming more important in data assimilation.
Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
  - Does not require prior knowledge of model error statistics.
Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
  - Does not require prior knowledge of model error statistics.
- Weak-constraint formulations a possible way forward - Requires specification of statistics.
Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
  - Does not require prior knowledge of model error statistics.
- Weak-constraint formulations a possible way forward - Requires specification of statistics.
- Different formulations lead to optimization problems with different sensitivities.